

Wi-Fi インバウンドカウンタによる 日本人・訪日外国人比率の推定手法

望月 祐洋¹・西田 純二²

¹非会員 (株) 社会システム総合研究所 (〒 550-0002 大阪府大阪市西区江戸堀 1-22-4)

E-mail: moma@jriss.jp

²フェロー会員 (株) 社会システム総合研究所 (〒 550-0002 大阪府大阪市西区江戸堀 1-22-4)

E-mail: nishida@jriss.jp

日本政府観光局の統計では 2018 年の訪日外客数は 3100 万人を越え、東京五輪・パラリンピック開催の 2020 年の目標として政府が掲げる年間訪日外国人旅行者数 4000 万人に達する勢いである。インバウンド客数の急増にともなうオーバーツーリズム等の課題に政策的に対処するには、その実態を適切に把握する必要がある。このために、アンケートベースのインバウンド調査や国土交通省が公開する訪日外国人流動データ (FF-Data) を利用して四半期別・年別に訪問地、国籍、利用交通機関等の属性に基づくクロス分析を実施する等の手段が存在する一方で、商店街や観光スポット等のより狭いエリアを対象にインバウンド客の現状を把握する簡便な手段が存在すれば政策選択の幅を広げられる。そこで本論文では、Wi-Fi パケットセンサをベースに試作した Wi-Fi インバウンドカウンタを用いた日本人と訪日外国人の比率の推定手法について報告する。

Key Words: インバウンド, 交通情報, 交通行動調査, Wi-Fi パケットセンシング, ICT

1. はじめに

日本政府観光局の統計では 2018 年の訪日外客数は 3100 万人を越え、東京五輪・パラリンピック開催の 2020 年の目標として政府が掲げる年間訪日外国人旅行者数 4000 万人に達する勢いである。インバウンド客数の急増にともなうオーバーツーリズム等の課題に政策的に対処するには、その実態を適切に把握する必要がある。このために、アンケートベースのインバウンド調査¹⁾や国土交通省が公開する訪日外国人流動データ (FF-Data)²⁾を利用して四半期別・年別に訪問地、国籍、利用交通機関等の属性に基づくクロス分析を実施する等の手段が存在する一方で、商店街や観光スポット等のより狭いエリアを対象にインバウンド観光客の現状を把握する簡便な手段が存在すれば政策選択の幅を広げられる。

我々は、Wi-Fi パケットセンサ³⁾をベースに観測エリア内の日本人・訪日外国人の比率を推定する機能拡張を施した Wi-Fi インバウンドカウンタを試作し、観光流動調査や交通行動調査などで訪日外国人の動向を分析する手段としての活用をめざしている。

本論文では、提案手法の前提、既存研究、Wi-Fi インバウンドカウンタのしくみを概説したうえで、予備実験として 2018 年 10 月 17 日 (水) に戎橋筋南とはり重前の 2 地点で実施した手計測とセンサ値との比較結果について報告し、最後にまとめと今後の課題について述べる。

2. 提案手法

(1) 前提

提案手法はスマートフォン等の Wi-Fi 対応機器がアクセスポイント (AP) を探すために定期的送信する管理パケット (プローブリクエスト) を収集する Wi-Fi パケットセンサの使用が前提となる。また、Wi-Fi パケットセンサの観測対象は Wi-Fi 機能を有効化している機器であり、Wi-Fi 機能を無効化している場合やそもそも Wi-Fi 対応機器を保有していない場合は観測されないため観測値と実数値とは一致しない。

(2) 既存研究

収集されたプローブリクエストのデータからスマートフォン利用者に関する何らかの属性情報を取得するアプローチとしては、一定期間 (例えば 1 週間) 収集したデータの中で一定日数 (例えば 3 日間) 以上観測される対象を住人、それ以外を訪問者とみなす方法⁴⁾、あるいは、Captive Portal 方式の Free Wi-Fi AP を Wi-Fi パケットセンサと組み合わせ、Free Wi-Fi へのオンライン登録の時点で氏名・国籍・性別といった属性情報を入力してもらうと同時に属性情報の提供の承認を得る手法⁵⁾が提案されている。

一方、プローブリクエストに含まれる SSID フィールドの値から属性情報を取得する手法が提案されている。SSID はそもそもは AP の識別名であり、自然言語

を用いて 0 から最大 32 バイト長の任意の名称を設定可能である。Wi-Fi 対応機器は過去に接続した AP のリスト (Preferred Network List; PNL) を保持しており、このリストに含まれる SSID をプローブリクエストの SSID フィールドに設定して送信することで、特定の SSID を持つ AP を迅速に発見できる。プローブリクエストはディレクティッドプローブ (directed probe) とブロードキャストプローブ (broadcast probe) の 2 種類に分類され、前者のプローブリクエストフレームには Wi-Fi 対応機器が過去に接続した AP の SSID が含まれる。これに対して後者のプローブリクエストフレームには SSID が含まれない。

この手法を採用する研究例としては、プローブリクエストに含まれるタイムスタンプ、送信元 (source) MAC アドレス、送信先 (destination) MAC アドレス、SSID (Service Set ID) フィールドの値を Wi-Fi フィンガープリントとして利用し、その類似度を計算することで利用者同士の関係を推定する SSID プロファイリングの研究が挙げられる⁶⁾。また、スマートフォン利用者同士のソーシャルネットワークを自動生成する研究⁷⁾では、SSID に使用されている自然言語が、利用者の国籍または利用者が長期滞在する国の使用言語に対応するという直感に基づき使用言語の自動検知が用いられている。さらに、世界中の Wi-Fi ホットスポットの SSID、MAC アドレス、GPS 座標情報をデータベース化した WiGLE 等のジオコーディングサービスを利用して SSID を位置情報に変換することで、大規模なイベントに参加した群衆の出身地を推定する研究なども行われている⁸⁾。

こうした研究の一方で、SSID として設定される文字列には、会社・大学・組織に紐づいた名称、参加した学術会議の情報、ホテル・レストラン・喫茶店・空港等の訪問地、“someone’s iPhone”など個人のアイデンティティに関わる情報が観察される⁶⁾ため、プローブリクエストとともに個人のプライバシーに関わる情報が拡散されるのを防ぐために、位置情報に基づいて Wi-Fi 対応機器の近隣にない AP へのプローブを抑制するといったプローブのしくみ自体を改善する研究⁹⁾も行われているなど、機微情報については注意深い取り扱いが必要である。

(3) Wi-Fi インバウンドカウンタ

我々は Wi-Fi パケットセンサの機能を拡張することで、日本人と訪日外国人の比率を算出する Wi-Fi インバウンドカウンタを試作した。Wi-Fi インバウンドカウンタでは、このためにプローブリクエストに含まれる SSID フィールドの値を利用しているが、その処理方法に特徴がある。既存の手法では、図 1 に示すように Wi-Fi パケットキャプチャ機器で収集した SSID を処理

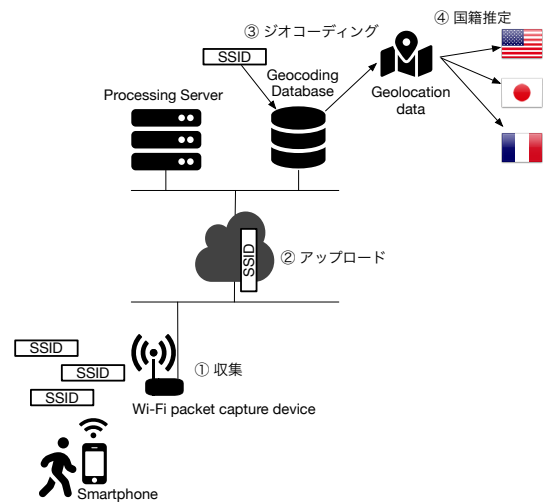
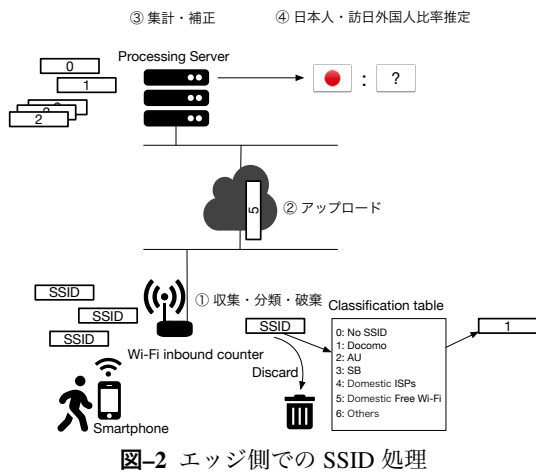


図-1 サーバ側での SSID 処理

サーバにアップロードして蓄積し、ジオコーディングデータベースに問い合わせることで位置情報に変換し、利用者の居住地または国籍を推定するというステップを踏むが、個人のプライバシーに関わり得る情報をサーバに蓄積するリスクが発生する。

これに対して提案手法では、図 2 に示すように、Wi-Fi インバウンドカウンタで収集した SSID をエッジ側で分類表に基づき分類番号に変換したうえで処理サーバにアップロードする。SSID の情報は分類番号への変換時点で即時破棄する。サーバ側では分類番号に基づいて集計・補正処理を行ったうえで日本人と訪日外国人の推定比率を算出する。具体的には、日本人の所有するスマートフォンは主に 3 大携帯キャリア、DOCOMO・SB・AU 等の通信会社の Wi-Fi に接続すると仮定し、これらに対応する分類番号のカウンタ数を日本人とみなす。

分類しきれなかったデータ (分類番号 0) については、処理サーバでは IEEE (Institute of Electrical and Electronics Engineers) が一元管理して製造ベンダに割り当てる OUI (Organizationally Unique Identifier) を国別に分類したテーブルを保持しており、OUI の値を 2 文字の国名コードに変換することでさらなる国別の分類・集計を試みる。ここで得られる国名コードはあくまで製造ベンダの所属国であり、スマートフォン利用者の国籍とみなすことはできないが、補正情報の一つとして利用する。OUI の検索・分類の過程で、製造ベンダが“Apple, Inc.”の割合が相当数にのぼるが、これを機械的に分類すると国名コードは US となり訪日外国人にカウントされてしまう。そこで Apple 製のデバイスについては国別の集計に含めず別途 Apple 製品として集計し、国別の Android・iOS のシェアの割合によって Apple 製品の集計値を按分することで補正を行っている。



(4) インバウンド率

インバウンドカウンタを用いて算出した訪日外国人の推定割合をインバウンド率と呼ぶ。インバウンド率は、1 時間毎に過去 1 時間のデータを利用して算出される。インバウンド率の変化の可視化例として 2018 年 10 月 31 日 (水) のハロウィン前後の戒橋北詰 (道頓堀) におけるインバウンド率の推移を図 3 に示す。

10 月 31 日の通勤時間帯には、日本人通勤者が増加するためインバウンド率は低下する (09:00 に 21.5%)。その後インバウンド率は 30% 台を推移するが、ハロウィンのために集まる日本人が深夜に向けて増加するため、ピーク時 (11 月 01 日の 01:00) のインバウンド率は 16.22% まで低下する。通常の推移としては (11 月 02 日) 16 時頃からインバウンド率が増加し 40% を超える。

なお、インバウンド率の算出に用いられるサンプルサイズは経時的に変化する点、および算出には一定以上のサンプルサイズが必要となる点に注意が必要である。

3. 予備実験

予備実験として 2018 年 10 月 17 日 (水) に戒橋筋南とはり重前の 2 地点で手計測を実施し、Wi-Fi インバウンドカウンタとの比較を行った。インバウンド率の推移を図 4 に、手計測とセンサ値の比較を表 3 に示す¹。

はり重前のインバウンド率は、朝の 6 時・7 時に低くなる一方、19 時以降に高くなり 20%~30% 程度となる。はり重前は自動車による通過を一部含む数値であるため、歩行者だけのインバウンド率はこの数値より 5%~10% 程度高いと考えられる。

戒橋筋南のインバウンド率は、昼間から夜間にかけて 11%~20% 程度と、はり重前よりも低くなっている。しかしながら、午前 02:00 から 05:00 にかけて日本人が

¹ センサ値は自動車交通を含むため、歩行者との比較で差が発生していると思われる。また、外国人率は目視調査であるため、観測者の感覚的判断に依存している。

表-1 はり重前インバウンドカウンタ算出値 (2018/10/17)

時刻	日本人推定件数	外国人推定件数	日本人割合 [%]	外国人割合 [%]
00:00	279	103	73.04	26.96
01:00	223	72	75.68	24.32
02:00	159	45	77.89	22.11
03:00	123	31	79.98	20.02
04:00	118	26	81.72	18.28
05:00	165	44	78.91	21.09
06:00	224	41	84.43	15.57
07:00	316	73	81.12	18.88
08:00	405	105	79.45	20.55
09:00	555	177	75.80	24.20
10:00	505	136	78.78	21.22
11:00	590	154	79.30	20.70
12:00	595	141	80.87	19.13
13:00	599	173	77.60	22.40
14:00	614	161	79.20	20.80
15:00	694	147	82.51	17.49
16:00	667	173	79.41	20.59
17:00	694	212	76.57	23.43
18:00	768	224	77.45	22.55
19:00	666	246	73.04	26.96
20:00	396	147	72.89	27.11
21:00	626	229	73.23	26.77
22:00	532	210	71.66	28.34
23:00	402	169	70.36	29.64

表-2 戒橋筋南インバウンドカウンタ算出値 (2018/10/17)

時刻	日本人推定件数	外国人推定件数	日本人割合 [%]	外国人割合 [%]
00:00	183	68	72.85	27.15
01:00	95	62	60.73	39.27
02:00	79	56	58.32	41.68
03:00	47	70	39.86	60.14
04:00	58	61	48.71	51.29
05:00	106	70	60.27	39.73
06:00	148	74	66.72	33.28
07:00	364	71	83.74	16.26
08:00	473	83	85.14	14.86
09:00	525	97	84.39	15.61
10:00	691	130	84.11	15.89
11:00	805	148	84.50	15.50
12:00	767	114	87.01	12.99
13:00	849	133	86.46	13.54
14:00	880	117	88.28	11.72
15:00	949	129	88.05	11.95
16:00	827	123	87.01	12.99
17:00	957	148	86.62	13.38
18:00	952	156	85.90	14.10
19:00	841	193	81.38	18.62
20:00	483	117	80.44	19.56
21:00	594	142	80.73	19.27
22:00	595	118	83.43	16.57
23:00	474	98	82.90	17.10

いない早朝の時間帯には、インバウンド率が異常に高く、外国人の流動が深夜・早朝も継続していることが見てとれる。この理由は、近隣のドン・キホーテ等の 24 時間営業の店舗に深夜から早朝まで外国人が来訪している様子を反映していると考えられる。

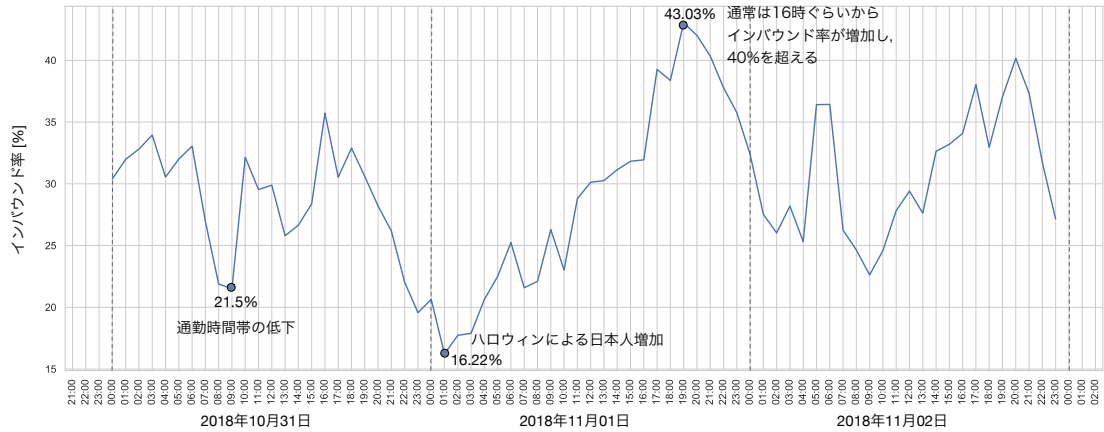


図-3 ハロウィン前後の戎橋筋北詰(道頓堀)におけるインバウンド率の推移

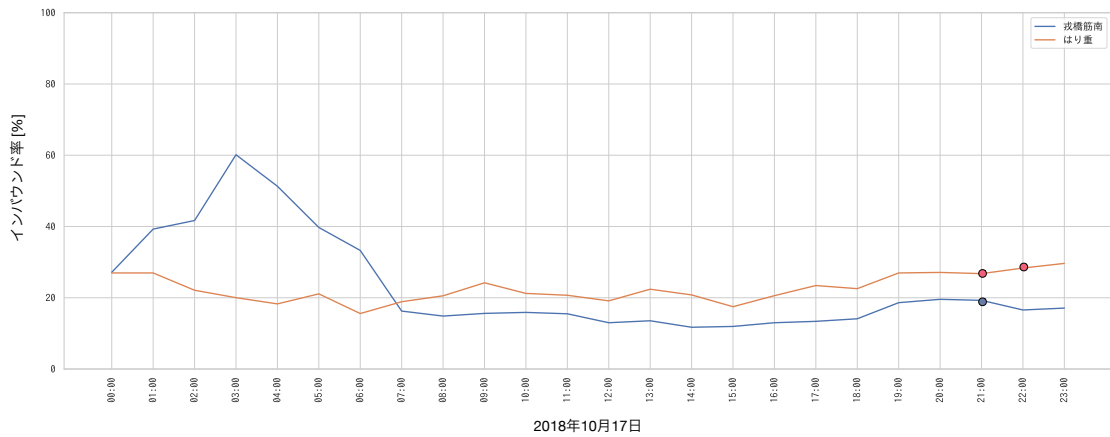


図-4 2018年10月17日(水)の戎橋南およびはり重前におけるインバウンド率の推移

表-3 2018年10月17日の手計測とセンサ値の比較

地点	戎橋筋南	はり重前	はり重前
時間	21:00-21:05	21:18-21:23	22:30-22:35
断面交通量	181	265	213
外国人	44	90	73
外国人の率	24.3%	34.0%	34.3%
センサ値	16.6%	28.3%	29.6%
計測の差	-7.7%pt	-5.7%pt	-4.7%pt

実験を実施して Wi-Fi インバウンドカウンタで算出したインバウンド率と手計測による観測値を比較したところ、センサ値の方が手計測の結果に比べて戎橋筋南では7.7%ポイント、はり重前では最大5.7%ポイント低い値を示した。この差はセンサ値が自動車交通を含むために生じたと考えられる。今後さまざまな実証実験を通じて、さらなる検証を進める予定である。

4. まとめと今後の課題

本論文では、Wi-Fi パケットセンサをベースに試作した Wi-Fi インバウンドカウンタを用いた日本人と訪日外国人の比率の推定手法について報告した。本手法の特徴は、Wi-Fi インバウンドカウンタで収集した SSID の情報をエッジ側で分類番号に変換すると同時に破棄することで、個人のプライバシーに関わる情報がサーバ側に蓄積されるリスクを回避している点にある。予備

参考文献

- 1) トリップアドバイザー: インバウンドレポート 2019, <https://tg.tripadvisor.jp/news/blog/inboundreport2019/>, 2019.
- 2) 国土交通省: FF-Data (訪日外国人流動データ), http://www.mlit.go.jp/sogoseisaku/soukou/sogoseisaku_soukou_fr_000022.html, 2017.
- 3) 望月祐洋, 上善恒雄, 西田純二, 中野秀男, 西尾信彦: Wi-Fi パケットセンサを利用した匿名人流解析システムの構築, 情報処理学会研究報告. MBL, [モバイルコンピューティングとユビキタス通信研究会研究報告], Vol.2014, No.45, pp.1-8, mar 2014.

- 4) 浅尾啓明, 西田純二, 安東直紀, 前田繁: Wi-Fi パケットセンサーによる交通常時観測システムの実用化, 土木計画学会第 55 回土木計画学研究発表会春大会, No.55, 2017.
- 5) 秋山豊和, 上東亜佑稀, 石原真太郎, 望月祐洋, 西田純二: Wi-Fi を用いた人流分析における課題とその解決に向けた Free Wi-Fi 認証機能の検討, JSPS 163rd Committee on Internet Technology (ITRC), November 2017.
- 6) Cunche, M., Kaafar, M. A., and Boreli, R.: Linking wireless devices using information contained in wi-fi probe requests, *Pervasive and Mobile Computing*, Vol.11, 04 2014.
- 7) Barbera, M. V., Epasto, A., Mei, A., Perta, V. C., and Stefa, J.: Signals from the crowd: Uncovering social relationships through smartphone probes, *Proceedings of the 2013 Conference on Internet Measurement Conference, IMC '13*, pp. 265–276, New York, NY, USA, 2013, ACM.
- 8) Di Luzio, A., Mei, A., and Stefa, J.: Mind your probes: De-anonymization of large crowds through smartphone wifi probe requests, *IEEE INFOCOM 2016 - The 35th Annual IEEE International Conference on Computer Communications*, pp. 1–9, April 2016.
- 9) Kim, Y. S., Tian, Y., Nguyen, L. T., and Tague, P.: Lapwin: Location-aided probing for protecting user privacy in wi-fi networks, *2014 IEEE Conference on Communications and Network Security*, pp. 427–435, Oct 2014.

(2019. 9. 27 受付)